

일반논문 (Regular Paper)

방송공학회논문지 제26권 제4호, 2021년 7월 (JBE Vol. 26, No. 4, July 2021)

<https://doi.org/10.5909/JBE.2021.26.4.453>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

인조 번호판을 이용한 자동차 번호인식 성능 향상 기법

장 일 식^{a)}, 박 구 만^{b)†}

Improved Method of License Plate Detection and Recognition using Synthetic Number Plate

Il-Sik Chang^{a)} and Gooman Park^{b)†}

요 약

자동차 번호인식을 위해선 수많은 번호판 데이터가 필요하다. 번호판 데이터는 과거의 번호판부터 최신의 번호판까지 균형 있는 데이터의 확보가 필요하다. 하지만 실제 과거의 번호판부터 최신의 번호판의 데이터를 획득하는데 어려움이 있다. 이러한 문제를 해결하기 위해 인조 번호판을 이용하여 자동차 번호판을 생성하여 딥러닝을 통한 번호판 인식 연구가 진행되고 있다. 하지만 인조 데이터는 실제 데이터와 차이가 존재하며, 이러한 문제를 해결하기 위해 다양한 데이터 증강 기법을 사용한다. 기존 데이터 증강 방식은 단순히 밝기, 회전, 어파인 변환, 블러, 노이즈등의 방법을 사용하였다. 본 논문에서는 데이터 증강 방법으로 인조데이터를 실제 데이터 스타일로 변환하는 스타일 변환 방법을 적용한다. 또한 실제 번호판 데이터는 원거리가 많고 어두운 경우 잡음이 많이 존재한다. 단순히 입력데이터를 가지고 문자를 인식할 경우 오인식의 가능성이 높다. 이러한 경우 문자인식 향상을 위해 본 논문에서는 문자인식을 위하여 화질개선 방법으로 DeblurGANv2 방법을 적용하여 번호판 인식 정확도를 높였다. 번호판 검출 및 번호판 번호인식을 위한 딥러닝의 방식은 YOLO-V5를 사용하였다. 인조 번호판 데이터 성능을 판단하기 위해 자체적으로 확보한 자동차 번호판을 수집하여 테스트 셋을 구성하였다. 스타일 변환을 적용하지 않은 번호판 검출이 0.614mAP를 기록하였다. 스타일 변환을 적용한 결과 번호판 검출 성능이 0.679mAP 기록하여 성능이 향상되었음을 확인하였다. 또한 번호판 문자인식에는 화질 개선을 하지 않은 검출 성공률은 0.872를 기록하였으며, 화질 개선 후 검출 성능이 0.915를 기록하여 성능 향상이 되었음을 확인 하였다.

Abstract

A lot of license plate data is required for car number recognition. License plate data needs to be balanced from past license plates to the latest license plates. However, it is difficult to obtain data from the actual past license plate to the latest ones. In order to solve this problem, a license plate recognition study through deep learning is being conducted by creating a synthetic license plates. Since the synthetic data have differences from real data, and various data augmentation techniques are used to solve these problems. Existing data augmentation simply used methods such as brightness, rotation, affine transformation, blur, and noise. In this paper, we apply a style transformation method that transforms synthetic data into real-world data styles with data augmentation methods. In addition, real license plate data are noisy when it is captured from a distance and under the dark environment. If we simply recognize characters with input data, chances of misrecognition are high. To improve character recognition, in this paper, we applied the DeblurGANv2 method as a quality improvement method for character recognition, increasing the accuracy of license plate recognition. The method of deep learning for license plate detection and license plate number recognition used YOLO-V5. To determine the performance of the synthetic license plate data, we construct a test set by collecting our own secured license plates. License plate detection without style conversion recorded 0.614 mAP. As a result of applying the style transformation, we confirm that the license plate detection performance was improved by recording 0.679mAP. In addition, the successful detection rate without image enhancement was 0.872, and the detection rate was 0.915 after image enhancement, confirming that the performance improved.

Keyword : Synthetic license plate, Data augmentation, Style transformation, DeblurGANv2, YOLO-V5

1. 서론

딥러닝의 지도 학습은 많은 양의 훈련 데이터와 훈련 데이터에 레이블을 동시에 학습시키는 방식으로 이루어진다. 데이터의 클래스가 불균형할 경우 불균형한 데이터 집합에서 많은 부분을 차지하는 다수 클래스와 적은 부분을 차지하는 소수 클래스의 샘플 개수를 균형 있는 데이터 셋으로 만든다. 어떤 클래스의 데이터 개수를 조절하는 가에 따라 언더샘플링 기법과 오버샘플링 기법으로 분류된다^[1]. 오버샘플링은 적은 클래스의 데이터를 랜덤하게 복제하여 수를 늘리는 것(Random Over-Sampling)이다. 그리고 언더샘플링은 적은 클래스의 데이터 크기에 맞도록 다수 클래스의 샘플들을 랜덤하게 제거하는 방식(Random Under-Sampling)이다. 번호판의 경우 과거의 번호판 데이터부터 최신의 번호판 데이터까지 다양한 종류가 존재한다. 다양한 종류의 번호판 데이터를 균형 있게 확보하는데 어려움이 있다. 이러한 클래스의 불균형 문제점을 해결 하고자 인조데이터를 가지고 번호판 데이터를 학습하는 연구^[2]가 진행 되고 있다. 인조 데이터를 가지고 실제 데이터와 비슷한 환경을 구축하여 높은 성능을 위하여 사용하는 방법으로 도메인 적응(Domain Adaptation)^[3] 방법이 있다. 인조 번호판 데이터와 실제 번호판 데이터간 도메인 차이에 의한 성능 하락을 보상하기 위해, 특정 도메인의 데이터로 학습된 모델을 다른 도메인에 적합하도록 변환하는 작업을 도메인 적응이라 한다. 하지만 본 논문에서는 실제 번호판의 데이터없이 학습을 진행하기 때문에 적용하지 않았다. 본 논문에서는 데

이터 증강 방식을 사용하여 실제 데이터와 비슷한 환경을 구축하였다. 데이터 증강 방법은 모델 학습에 사용되는 데이터의 다양한 변이를 배움으로써, 학습 모델의 과적합(Over-fitting)을 방지하기 위해서 유용하게 사용된다. 데이터 증강 기법을 크게 분류하면, 영상처리 기반의 증강 방법과 적대적 생성 네트워크(Generative Adversarial Network)^[4] 기반의 증강 방법이 있다. 영상처리 기반의 증강 방법에는 영상의 기하학적(Geometric) 변환, 컬러 공간 변환, 커널 필터, 랜덤 지움, 이미지 혼합 등이^[5] 있다. 하지만 이러한 이미지 밝기 혹은 공간적 정보를 변형하는 등의 데이터 증강 방법은 변형 전 원본 이미지의 속성과 많은 부분 닮아있다. 따라서 이미지가 가진 속성들의 다양한 변이를 학습하지 못해 큰 성능 향상을 이루지 못하는 경우가 많다. 이러한 한계를 극복하기 위해 적대적 생성 네트워크 기반 데이터 증강 방법이 다양한 연구에서 적용되고 있다. 또한 최근에는 학습 이미지의 일부를 잘라 내거나 이미지들을 섞는 방식^{[6][7][8][9]}으로 간단하면서도 학습 성능을 끌어 올릴 수 있는 데이터 증강 방법들도 소개되고 있다. 본 논문에서는 정형화된 번호판 데이터를 생성하기 위해 인조 번호판 데이터 생성 프로그램을 개발하였다. 기존 인조 번호판 생성 프로그램은 번호판에 잡음, 크기 변환 등의 기법을 임의로 적용하여 다양한 형태의 인조 번호판을 생성하였다. 인공 번호판에 실제 번호판의 스타일 변환과정을 거쳐 단순 데이터 증강으로 학습된 결과 대비 인조 번호판의 검출 성능을 향상 시켰다. 또한 검출된 번호판의 해상도가 낮거나 잡음을 많이 가지고 있어 번호판의 문자판별의 정확도가 낮아지는 문제가 있다. 본 논문에서는 검출된 해상도가 낮은 번호판을 적대적 생성 네트워크인 DeblurGANv2^[10] 방법을 사용하여 이미지 향상 후 문자를 인식하여 성능의 정확도를 향상 시켰다.

II. 본론

1. 시스템 구성

그림 1은 본 논문에서 제안하는 인공번호판 데이터를 이

a) 서울과학기술대학교 나노IT디자인융합대학원 정보통신미디어공학전공(Dept. of Information Technology and Media Engineering, The graduate School of Nano IT Design Fusion, Seoul National University of Science and Technology)

b) 서울과학기술대학교 전자IT미디어공학과(Dept. of Electronic IT Media Engineering, Seoul National University of Science and Technology)

‡ Corresponding Author : 박구만(Goo-man Park)

E-mail: gmpark@seoul.ac.kr

Tel: +82-2-970-6425

ORCID: <http://orcid.org/0000-0002-7055-5568>

※ This research was supported by R&BD Program through the INNOPOLIS funded by Ministry of Science and ICT (2020-IT-RD-0232).

· Manuscript received May 28, 2021; Revised July 1, 2021; Accepted July, 1, 2021.

용한 훈련 파트와 추론 파트로 표현하였다.

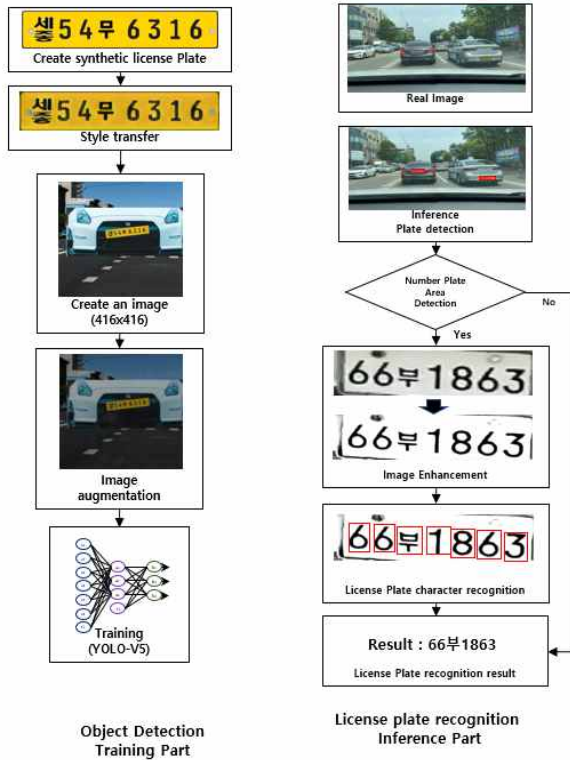


그림 1. 시스템 구성도
 Fig. 1. Block diagram of system

훈련 파트에서는 인공 데이터 생성을 통해 획득한 데이터를 실제 번호판 스타일로 변환해 준다. YOLO-V5^[11] 학습을 위해서는 416x416 크기로 주변 환경데이터 및 자동차 앞부분을 추가하여 실제 자동차의 이미지로 변환한다. 자동차 앞부분은 다수의 자동차 앞부분을 수집하여 랜덤하게 선정한다. 416x416 크기의 이미지는 다양한 환경에 일반화가 가능하도록 데이터 증강 방법을 수행한 후 학습을 진행한다. 학습은 번호판 검출과 번호판 글자 인식 부분으로 구분하여 2번의 학습을 진행한다. 추론 파트(Inference Part)에서는 학습된 결과를 이용하여 실제 이미지(Real Image)를 입력으로 하여 자동차 번호판을 검출한다. 검출된 번호판 영역은 DeblurGANv2로 학습된 이미지 개선 단계를 거쳐 이미지를 향상 후 번호판의 글씨를 인식한 후 최종 결과값을 출력한다.

2. 관련연구

2.1 딥러닝 기반 객체 검출

딥러닝 기반 객체 검출은 크게 1단계 검출 방식과 2단계 검출 방식으로 나눌 수 있다. 2단계 검출 방식은 객체의 후보 영역(Region Proposal)을 검출 후 객체의 후보 영역에 대한 객체 분류 및 객체의 위치를 추정하는 방식이다. 1단계 검출 방식은 두 가지 단계를 한 번에 처리하는 방식을 말한다. 1단계 검출 방식은 비교적 빠르지만 정확도가 낮고 2단계 검출 방식은 비교적 느리지만 정확도가 높은 특징이 있다. 2단계 검출 방식의 대표적인 방법으로 R-CNN^[12], Fast R-CNN^[13], Faster R-CNN^[14] 등이 있고 1단계 검출 방식으로는 SSD 계열^[15]과 YOLO(You Look Only Once)^{[16][17][18][19][11]} 계열이 있다. 본 논문에서는 객체 검출을 위해 실시간 처리가 가능하고, 가장 최근에 발표한 Yolo-V5를 사용하였다. Yolo-V5는 2020년 6월에 Glenn Jocher가 발표한 파이토치(Pytorch)를 사용하는 객체 검출 방법이다. Yolo-V5의 모델은 4가지로 나눌 수 있는데 표 1과 같다.

표 1. YOLO-V5 모델
 Table 1. YOLO-V5 Model

Model	size	Speed v100(ms)	params (M)	FLOPS 640(B)
YOLOv5s	640	2.0	7.3	17.0
YOLOv5m	640	2.7	21.4	51.3
YOLOv5l	640	3.8	47.0	115.4
YOLOv5x	640	6.1	87.7	218.8

본 논문에서는 m 모델과 l 모델을 사용하였다.

2.2 스타일 변환

스타일 변환의 대표적인 방법은 Image Style Transfer Using Convolutional Neural Networks^[20] 방법이 있다. 해당 방식은 CNN을 활용해 이미지의 스타일을 전송하는 방법을 제안하였다. 스타일 변환을 위해 사용되는 손실은 식 (1)과 같이 콘텐츠 손실과 스타일 손실로 구분한다.

$$L_{total}(\vec{p}, \vec{a}, \vec{x}) = \alpha L_{content}(\vec{p}, \vec{x}) + \beta L_{style}(\vec{a}, \vec{x}) \quad (1)$$

전체 구조는 그림 2와 같다.

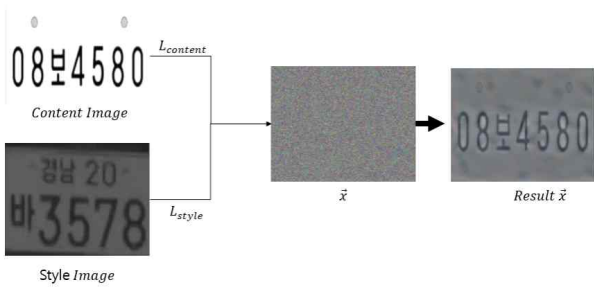


그림 2. Image Style Transfer Using CNN 구성도
Fig. 2. Block diagram of Image Style Transfer Using CNN

해당 방법은 스타일 변환을 위해 매번 학습이 필요하기 때문에 시간이 오래 걸리는 문제가 있다. 이러한 문제를 개선한 AdaIN Style Transfer^[21] 방식은 자신이 원하는 임의의 스타일 이미지로부터 스타일 정보를 가져와서 실시간으로 빠르게 스타일 변환이 가능하다. 그림 3은 AdaIN Style Transfer 전체 구성을 나타낸다. AdaIn는 식(2)와 같이 특징 공간 내에서 특징 통계량(채널 단위의 평균, 분산)을 전달함으로써 스타일 변환 역할을 수행한다.

$$AdaIn(x, y) = \sigma(y) \left(\frac{x - \mu(x)}{\sigma(x)} \right) + \mu(y) \quad (2)$$

x : 콘텐츠 이미지, y : 스타일 이미지

AdaIn을 이용하여 다른 원하는 이미지에서 스타일 정보를 가져와 식(3)과 같이 적용이 가능하다. 본 논문에서는 가상의 인조 번호판에 실제 번호판 스타일을 적용하였다. α (0.0 ~ 1.0) 값을 통하여 동적으로 스타일 반영 정도를 조절 가능하다.

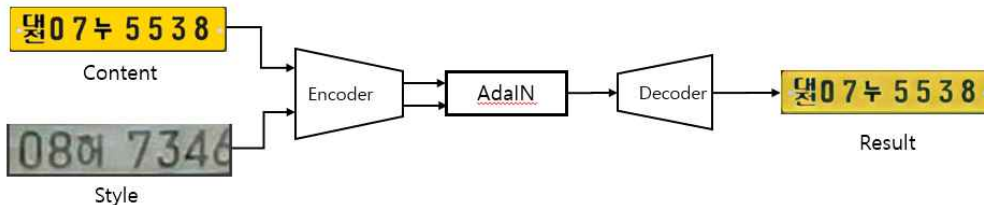


그림 3. AdaIN Style Transfer 구성도
Fig. 3. Block diagram of AdaIN Style Transfer

$$T(c, s, \alpha) = g((1 - \alpha)f(c) + \alpha AdaIN(f(c), f(s))) \quad (3)$$

$AdaIN(f(c), f(s))$: AdaIN 스타일 변환

$f(c), f(s)$: encoder에 의해 압축된 latent vector

그림 4는 α 값에 따른 번호판 스타일 변환이 적용된 결과를 나타낸다.

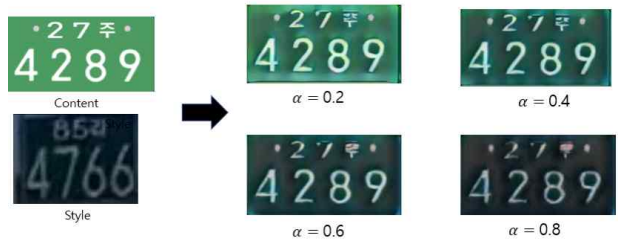


그림 4. α 값에 따른 스타일 변환 결과
Fig. 4. Style Transfer result of α value

α 값이 너무 큰 경우 번호판 데이터가 심하게 바뀌는 문제가 발생하였다. 본 논문에서는 이러한 문제점을 고려하여 α 값을 0~0.5 값으로 랜덤하게 생성하여 스타일 변환을 한다. 다양한 실제 번호판을 수집하여 스타일 변환을 할 경우 랜덤하게 실제 번호판 선택하여 실제 번호판 스타일로 변환을 함으로써 번호판의 검출 성능에 향상을 주었다. 스타일 변환을 위해 쌍을 이루지 않은(Unpaired) 데이터로 학습하는 cycleGAN^[22]을 적용하여 스타일 변환을 수행하였으나 좋은 성능의 결과가 나오지 않았다.

2.3 인조 번호판 생성

본 논문의 인조 번호판 생성방법은 기존 “3차원 회전을 이용한 인조 번호판 생성기의 번호판 인식 성능 비교”^[23] 방법을 기반으로 3차원 인조 번호판을 생성하였다. 인조 번

호판 생성 순서는 번호판의 종류를 랜덤하게 선택하고, 번호판 배경의 규격에 맞는 숫자와 한글 이미지를 선택한다. 숫자와 한글은 번호판 규격에 맞게 랜덤하게 선택된다. 심볼 및 볼트는 규격에 맞게 선택 된다. 선택된 번호판, 숫자, 한글, 볼트 또는 심볼 이미지가 조합하여 번호판의 이미지를 만들게 된다. 인조 번호판 생성은 그림 5와 같다.

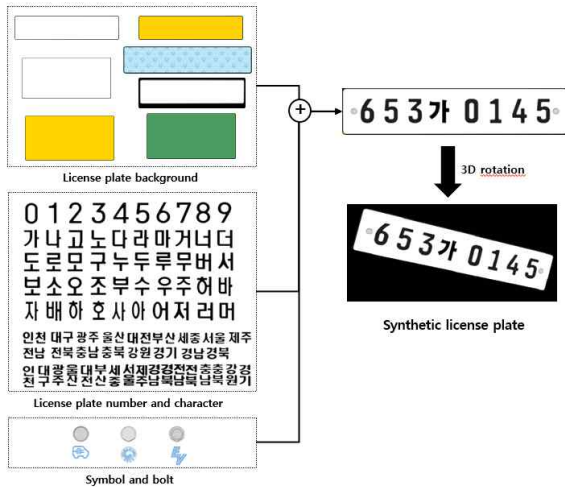


그림 5. 인조 번호판 생성
 Fig. 5. Create an synthetic license plate

2.4 데이터 증강

생성된 인조 번호판 이미지는 스타일 변환을 거쳐 실제 데이터와 유사한 환경을 만들기 위해 표 2에 정리된 내용과 같이 확률적으로 번호판을 생성한다. Fog의 경우 Auto-mold Road Augmentation Library^[24]를 사용하였다.

표 2. 데이터 증강 파라미터 정보
 Table 2. Data augmentation parameter information

Method	Parameter
Gray	Probability : 20%
Bright	Min factor : 0.6 Max factor : 1
Noise	Gaussian Probability : 67%
Fog	Coeff 0.08 ~ 0.4
Blur	Kernel : 3x3 Probability : 50%

다양한 데이터 증강 방식을 거쳐 최종 학습데이터로 사용한다. 학습데이터가 만들어지는 순서는 먼저 인조 번호판을 생성한 후 스타일 변환을 거쳐 번호판을 생성한다. 학습을 위한 크기인 416x416 크기의 배경으로 실제 배경 이미지를 사용하고, 차량 전면은 실제 자동차의 이미지를 사용하여 실제 환경과 유사하도록 구성 하였다. 그림 6은 실제 차량 전면에 사용한 이미지를 나타낸다.



그림 6. 차량 전면 이미지
 Fig. 6. Vehicle front image

배경과 차량 전면 그리고 번호판을 합성하여 만들어진 이미지는 데이터 증강 과정을 거쳐 최종 학습데이터로 생성된다. 그림 7은 가상 데이터 생성부터 데이터 증강까지의 과정을 나타낸 것이다.

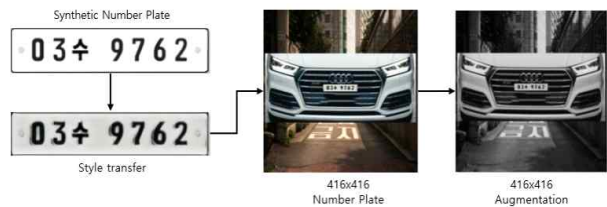


그림 7. 가상 데이터 생성부터 데이터 증강 과정
 Fig. 7. From synthetic data creation to data augmentation process

2.5 문자 검출 및 화질 개선

문자 검출을 위해서 총 85개의 클래스로 구분하여 학습을 진행한다. 본 논문에서는 번호판 검출단계와 검출된 번호판 영역에서 자동차 번호 문자 검출 단계로 구분하여 이루어진다. 실제 번호판 데이터는 야간이나 원거리 혹은 화질 저하로 인하여 문자 검출 성능이 떨어지게 된다. 특히 야간의 경우 번호판의 배경과 번호판 내부 글자와의 화소 값 차이가 적게 나기 때문에 번호 인식에 오류가 발생하였다. 이러한 문제점을 해결하기 위해 초해상도(Super-

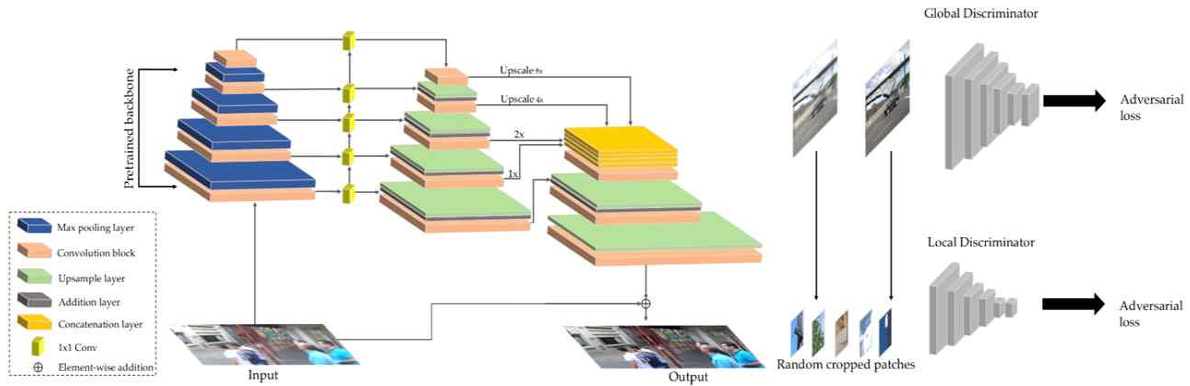


그림 8. DeblurGANv2 구조
Fig. 8. Structure of DeblurGANv2

Resolution) 알고리즘으로 SR-GAN^[25], EDSR^[26], ESR-GAN^[27] 등의 방법이 연구되고 있다. 하지만 실제 초해상화 방법으로 실험한 결과를 분석했을 때 결과 인조 데이터에 대한 좋은 결과가 나오지만 실제 이미지가 입력되었을 경우 원하는 결과가 나오지 않았다. 원인은 스타일 변환된 가상 데이터를 가지고 학습을 하였기 때문에 실제 데이터와 어느 정도 차이가 나는 것으로 보인다. 본 논문에서는 위의 문제점에 좋은 성능을 나타내는 방법으로 블러한 영상을 개선하는 방법을 사용하면 성능 향상이 될 것으로 판단하여 DeblurGAN^[28], DeblurGANv2 방법을 사용하였다. 실험 결과 성능 향상이 높게 나오는 DeblurGANv2를 사용하여 이미지의 화질을 개선하였다. DeblurGANv2는 적대적 신경망 네트워크(GAN)를 사용하며, 이중 스케일 판별기가 있는 상대론적 조건부 GAN을 기반으로 한다. 핵심 구성 요소로는 FPN(Feature Pyramid Network)을 디블러링(deblurring)에 처음 사용하고 백본으로는 Inception-ResNet-v2^[29], MobileNetV2^[30]를 사용한다. 본 논문에서는 백본으로 MobileNetV2를 사용하였다. 그림 8은 DeblurGANv2의 구조를 나타낸다.

3. 실험

본 논문에서는 자동차 번호판 및 자동차 번호판 글자 인식을 위하여 YOLO-V5를 사용하여 학습하였다. 그림 9는 학습에 사용된 자동차 번호판 이미지와 번호판 문자에 대한 이미지를 나타낸다.



그림 9. 학습에 사용된 번호판 이미지와 번호판 문자 이미지
Fig. 9. License plate image and licence plate character image used for learning

붉은 색 박스는 Ground Truth를 나타낸다. 정상적인 번호판은 숫자가 많고 한글이 적기 때문에 클래스 불균형이 심하다. 이러한 문제점을 개선하기 위하여 번호판 글씨 인식을 위해 번호판에 한글을 추가로 넣음으로써 클래스 간 불균형을 개선하였다. 번호판 검출을 위하여 50만장의 데이터를 사용하였다. 학습은 90%인 45만장은 학습데이터로 사용하였고, 5만장은 검증데이터로 사용하였다. 구축한 데이터 셋을 기준으로 학습 모델은 YOLOv5m, YOLOv5l을

이용하였으며 이미지의 크기는 416x146으로 설정하였다. 에폭은 총 15회로 하였다. 번호판 검출 테스트는 실제 도로에서 촬영한 영상을 가지고 3가지의 도로 영상으로 테스트 하였다. 화면의 크기는 1920x1080 이며 번호판의 크기가 25 픽셀 보다 작은 경우 무시하였다. 표 3은 번호판 검출에 대한 결과를 나타낸다.

표 3. 번호판 검출 결과
 Table 3. Result of plate number detection

Testset \ mAP	Style, m	Style, l	m	l
1	0.677	0.614	0.545	0.607
2	0.519	0.498	0.464	0.410
3	0.678	0.577	0.512	0.453

Style은 스타일이 적용된 데이터로 학습된 모델을 의미하고, m, l 은 YOLO V5 m, l 모델을 의미한다. 스타일이 적용된 m 모델로 학습하였을 경우 가장 좋은 성능이 나올 수 있다. 그림 10은 스타일 변환 적용할 경우와 하지 않았을 경우의 결과를 나타낸다.

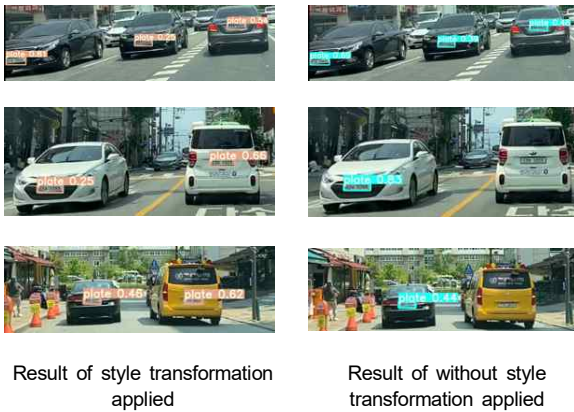


그림 10. 스타일 변환 적용 결과
 Fig. 10. Results of applying style transformation

스타일이 적용된 경우 번호판 검출이 성공하지만 스타일이 적용되지 않은 경우 번호판 검출이 실패하는 것을 볼 수 있다. 본 논문에서는 인조 번호판 이미지를 실제 번호판처럼 변환하기 위해 cycle-GAN 방법을 사용하여 실험하였

다. 결과는 그림 11과 같다.



그림 11. 인조 이미지에 cycle-gan 적용 결과
 Fig. 11. Results of applying cycle-gan to synthetic image

그림 11과 같이 cycle-GAN의 경우 인조 번호판 데이터와 실제 번호판 데이터를 가지고 학습한 결과 원하는 결과가 나오지 않았다. deblurGANv2는 384x96 크기로 10만장의 이미지를 이용하여 학습하였다. 학습데이터는 쌍으로 구성되며, 생성된 인조데이터와 인조 번호판에 실제 이미지의 스타일을 변경하고 데이터 증강 방식이 적용된 이미지를 쌍으로 사용하였다. 총 학습데이터는 10만장을 사용하였고, 10만장 중 실제 학습데이터는 95,000장으로 진행하였으며, 학습시 컬러를 사용할 경우 어두운 밤의 사진이나 간혹 실제 색상과 다르게 나오는 현상이 발생하여 회색(Gray) 이미지로 변경하여 학습을 진행하였다. 회색 이미지로 학습을 진행할 경우 원본의 컬러를 유지하고 번호판의 글자만 개선되었다. 그림 12는 데이터 증강을 통해 생성된 인조 번호판에 deblurGANv2 적용한 결과를 나타낸 그림이다.

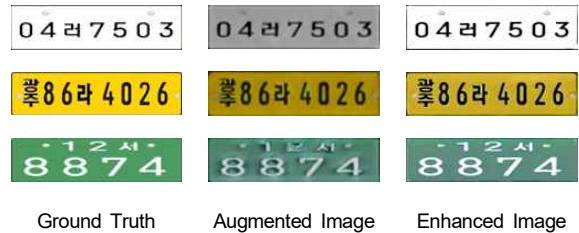


그림 12. 인조 이미지에 deblurGANv2 적용 결과
 Fig. 12. Results of applying deblurGANv2 to synthetic image

그림 13은 데이터 증강을 통해 생성된 실제 번호판에 deblurGANv2 적용한 결과를 나타낸 그림이다.



그림 13. 실제 이미지에 deblurGANv2 적용 결과
Fig. 13. Results of applying deblurGANv2 to real image

본 논문에서는 초 해상도 기법으로 ESR-GAN, EDSR 방법을 테스트 하였다. 결과는 그림14와 같다.

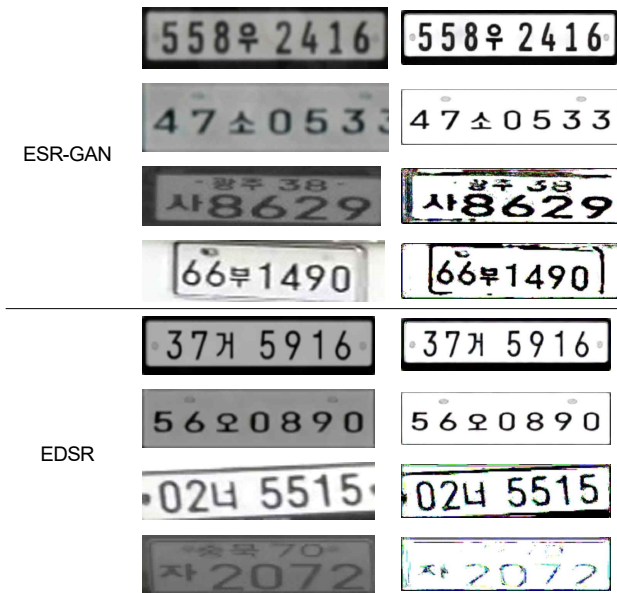


그림 14. 초해상도 방법의 결과
Fig. 14. Result of the super-resolution method

그림 14의 각각의 방법에서 위 2개는 인조 번호판에 대한 결과이고, 아래 2개는 실제 번호판의 결과이다. 인조 번호판의 경우엔 결과가 잘 나오지만 실제 번호판의 결과는 좋은 성능이 나오지 못했다. 이미지 개선을 위해 본 논문에서는 deblurGAN^[30], DeblurGANv2에 대하여 학습을 진행하였고, 테스트는 학습에 사용하지 않는 5000장에 대해

Ground truth와 개선된 이미지 간의 이미지 품질 측정 지표로 많이 사용하는 PSNR(Peak Signal-to-noise ratio), SSIM (Structural Similarity)를 사용하였다. 결과는 표 4와 같다. 본 논문에서는 좀 더 좋은 결과가 나온 DeblurGANv2를 사용하였다.

표 4. PSNR, SSIM 결과
Table 4. Result of PSNR and SSIM

	PSNR	SSIM
DeblurGAN	14.69	0.71
DeblurGANv2	18.67	0.79

번호판 글자 인식을 위해 총 60만장의 데이터를 사용하였다. 구축한 데이터 셋을 기준으로 모델은 YOLOv5m을 이용하였으며 에폭은 20으로하였다. 이미지의 크기는 416x146으로 설정하였다. 테스트시 사용한 총 번호판은 635장을 사용하였고 숫자 번호는 3818자 한글 문자는 720자를 사용하였다. 그림 15는 테스트 이미지에 Deblur GANv2가 적용여부에 따라 실제 번호 검출 결과를 나타낸다.



그림 15. DeblurGANv2 적용 결과
Fig. 15. Results of applying DeblurGANv2

붉은 색의 박스는 실제 번호판의 글자를 검출한 결과를

의미한다. 번호판의 화질이 좋을 경우에는 차이가 거의 없지만 번호판의 데이터가 야간의 경우 번호판 배경의 화소 값과 번호판의 번호 값의 차이가 크지 않을 경우 번호 인식이 잘 되지 못함을 그림15에서 확인 할 수 있었다. 번호판의 숫자 데이터와 한글 데이터 간 불균형이 심하기 때문에 mAP 뿐 아니라 글자 인식의 정확도를 같이 측정 하였다. 실제 테스트할 경우에는 실제 번호판을 사용하기 때문에 한글에 대한 클래스가 적게 나오게 된다. 글자 인식 정확도는 Ground Truth 데이터와 가장 큰 IOU를 기준으로 하여, 해당 위치에서 검출된 글자의 값과 Ground Truth의 실제 값과 동일한 결과일 경우 정상 인식으로 판별하였다. 이미지 개선 적용 후 번호판 글자 인식 결과는 표 5와 같다.

표 5. 이미지 개선 적용 후 번호판 글자 인식 결과
 Table 5. License plate character recognition result after applying image enhancement

	DeblurGANv2 Applied	DeblurGANv2 Not Applied
mAP	0.64	0.62
Total Accuracy	0.915	0.872
Number Accuracy	0.941	0.898
Character Accuracy	0.781	0.732
Plate Number Accuracy	0.737	0.732

III. 결 론

본 논문에서는 인조번호판을 이용하여 차량 번호판 인식 시스템을 제안하였다. 인조 번호판 생성을 통해 학습에 필요한 데이터를 확보 하였다. 인조 번호판을 사용함으로써 다양한 번호판의 종류를 균형 있게 생성이 가능하였다. 인조 번호판에 대한 데이터 증강 방법으로 다양한 방식이 있었지만 본 논문에서는 실제 번호판의 스타일 변환을 적용함으로써 실제 번호판 검출에 성능 향상이 됨을 확인 하였다. 번호판 글씨 인식을 위하여 번호판 검출과 동일한 방식을 사용하지만 번호판의 문자의 경우 숫자와 한글 등 총 85개의 클래스를 구분해야 하는 문제와 번호판 영역보다

작은 영역을 나타냄으로 정확한 검출에 어려움이 있다. 또한 실제 환경에서는 야간이나 원거리 등의 다양한 환경에서 잡음이나 왜곡에 의한 검출의 어려움이 있다. 이러한 문제를 개선하기 위해 본 논문에서는 학습 시 번호판에 한글 데이터를 추가로 포함시켰고, 이미지 화질 개선을 위해 DeblurGANv2 방법을 사용하였다. 화질이 개선된 번호판에서 번호 글씨 인식의 정확도가 향상됨을 확인 하였다. 본 논문에서 실험한 결과를 바탕으로, 추후에는 실제 자동차 영상에서 번호판 영역을 검출하여 인조 번호판을 적용하는 방법 및 인조 번호판을 실제 스타일로 변환하는 연구와 번호판의 화질 개선에 대한 연구도 지속할 예정이다.

참 고 문 헌 (References)

- [1] J. Van Hulse, T. M. Khoshgoftaar, A. Napolitano, "Experimental perspectives on learning from imbalanced data," in Proceedings of the ACM International Conference on Machine Learning, New York, pp.935-942, 2007.
- [2] Jae-Hyeon Lee, Sung-Man Cho, Seung-Ju Lee, Cheong-Hwa Kim, Goo-Man Park. "License Plate Recognition System Using Synthetic Data," pp.107 - 115, 2020, doi:10.5573/ieie.2020.57.1.107
- [3] Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, et al. "Domain-Adversarial Training of Neural Networks," Journal of Machine Learning Research vol. 17, pp.1-35, 2016
- [4] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. "Generative adversarial nets," Adv Neural Inf Process Syst. 2014.
- [5] Connor Shorten, Taghi M, Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," Journal of Big Data 2019.
- [6] Terrance Devries, Graham W, Taylor, "Improved regularization of convolutional neural networks with Cutout," arXiv preprint arXiv:1708.04552, 2017.
- [7] Hongyu Guo, Yongyi Mao, and Richong Zhang, "Mixup as locally linear out-of-manifold regularization," In AAAI, 2019.
- [8] Sangdoon Yun, Dongyoon Han, Seong Joon Oh, Sanghyuk Chun, Junsuk Choe, and Youngjoon Yoo, "Cutmix: Regularization strategy to train strong classifiers with localizable features," ICCV, 2019.
- [9] Dan Hendrycks, Norman Mu, Ekin Dogus Cubuk, Barret Zoph, Justin Gilmer, Balaji Lakshminarayanan. "AugMix: A Simple Data Processing Method to Improve Robustness and Uncertainty," ICLR, 2020.
- [10] Orest Kupyn, Tetiana Martyniuk, Junru Wu, Zhangyang Wang, "DeblurGAN-v2: Deblurring (Orders-of-Magnitude) Faster and Better," ICCV, 2019.
- [11] Glenn Jocher. <https://github.com/ultralytics/yolov5>
- [12] Ross Girshick, Jeff Donahue, Trevor Darrell, Jitendra Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," arXiv: 1311.2524v5, Oct 2014.

- [13] Ross Girshick, "Fast R-CNN," arXiv:1504.08083v2, Sep 2015.
- [14] Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," arXiv: 1506.01497v3, Jan 2016.
- [15] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, Alexander C. Berg, "SSD: Single Shot MultiBox Detector," arXiv:1512.02325v5, Dec 2016.
- [16] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," arXiv:1506.02640v5, May 2016.
- [17] Joseph Redmon, XNOR.ai, "YOLO9000: Better, Faster, Stronger," CVPR, 2017.
- [18] Joseph Redmon, Ali Farhadi, "YOLOv3: An Incremental Improvement. arXiv:1804.02767, Apr 2018.
- [19] Alexey Bochkovskiy, Chien-Yao Wang, Hong-Yuan Mark Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," arXiv:2004.10934, Apr 2020.
- [20] Leon A. Gatys, Alexander S. Ecker, Matthias Bethge, "Image Style Transfer Using Convolutional Neural Networks," CVPR. 2016.
- [21] Xun Huang Serge Belongie, "Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization," ICCV. 2017.
- [22] Jun-Yan Zhu, Taesung Park, Phillip Isola, Alexei A. Efros, "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks," ICCV. 2017.
- [23] Lee, Yu-Jin, Kim, Sang-Joon, Park, Gyeong-Moo, Park, Goo-Man, "Comparison of number plate recognition performance of Synthetic number plate generator using 2D and 3D rotation," The Korean Society Of Broad Engineers, pp.141-144, 2020.
- [24] Ujjwal Saxena. <https://github.com/UjjwalSaxena/Automold-Road-Augmentation-Library>
- [25] Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, Wenzhe Shi, "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network," CVPR. 2017.
- [26] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, Kyoung Mu Lee, "Enhanced Deep Residual Networks for Single Image Super-Resolution," CVPR. 2017.
- [27] Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Yu Qiao, Chen Change Loy, "ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks," ECCV. 2018.
- [28] Orest Kupyn, Volodymyr Budzan, et al. "DeblurGAN: Blind Motion Deblurring Using Conditional Adversarial Networks," CVPR. 2018.
- [29] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, "Inception-v4, inception-resnet and the impact of residual connections on learning," In: Thirty-First AAAI Conference on Artificial Intelligence. 2017.
- [30] Mark Sandler, Andrew Howard, Menglong Zhu, et al. "MobileNetV2: Inverted Residuals and Linear Bottlenecks," CVPR. 2018.

— 저 자 소 개 —



장 일 식

- 2011년 2월 : 서울과학기술대학교 NID융합기술대학원 석사
- 2020년 3월 ~ 현재 : 서울과학기술대학교 지능형미디어연구센터 책임연구원
- 2020년 9월 ~ 현재 : 서울과학기술대학교 나노IT디자인융합대학원 정보통신미디어공학전공 박사과정
- ORCID : <https://orcid.org/0000-0003-0822-9857>
- 주관심분야 : 컴퓨터비전, 딥러닝



박 구 만

- 1984년 2월 : 한국항공대학교 전자공학과 공학사
- 1986년 2월 : 연세대학교 전자공학과 공학석사
- 1991년 2월 : 연세대학교 전자공학과 공학박사
- 1991년 3월 ~ 1996년 9월 : 삼성전자 신호처리연구소 선임연구원
- 2016년 1월 ~ 2017년 12월 : 서울과학기술대학교 나노IT디자인융합대학원 원장
- 1999년 8월 ~ 현재 : 서울과학기술대학교 전자IT미디어공학과 교수
- 2006년 1월 ~ 2007년 8월 : Georgia Institute of Technology Dept.of Electrical and Computer Engineering, Visiting Scholar
- ORCID : <https://orcid.org/0000-0002-7055-5568>
- 주관심분야 : 컴퓨터비전, 지능형실감미디어